

ANÁLISE DE EFICIÊNCIA E COMPLEXIDADE DA CODIFICAÇÃO DA PROFUNDIDADE DE VÍDEOS DIGITAIS SEGUNDO O PADRÃO 3D-HEVC

GIOVANNI AVILA; RUHAN CONCEIÇÃO; MARCELO PORTO; BRUNO ZATT;
LUCIANO AGOSTINI; GUILHERME CORREA

Universidade Federal de Pelotas - Grupo de Arquiteturas e Circuitos Integrados
{gdadavila, radconceicao, porto, zatt, agostini, gcorrea}@inf.ufpel.edu.br

1. INTRODUÇÃO

Frente os recentes avanços nas tecnologias de vídeos digitais e sua crescente demanda, principalmente devido à popularização de serviços de *streaming* e as diversas aplicações de vídeos de alta resolução em dispositivos portáteis, faz-se cada vez mais necessário o uso de práticas para melhor manipular e transmitir esses dados. Tendo em vista este cenário e o fato de que um vídeo desprovido de qualquer tipo de compressão possui uma quantidade imensa de informações, que tornam seu armazenamento e manipulação muito difíceis, o uso de codificadores e decodificadores torna-se praticamente imprescindível.

Ao longo dos anos, diversos padrões de codificação foram estabelecidos e atualmente o *High Efficiency Video Coding* (HEVC) (JCT-VC, 2013) é considerado o estado-da-arte para vídeos 2D. Este codificador intensifica ainda mais a taxa compressão de vídeo em relação aos seus antecessores, mantendo a mesma qualidade de imagem, porém o HEVC faz isso ao custo de alta complexidade computacional, chegando a um aumento de quase cinco vezes em relação ao padrão anteriormente estabelecido.

Além do avanço no uso de vídeos digitais, também as aplicações de vídeos 3D apresentam uma popularização nos últimos anos. Como respaldo a esta situação, foi desenvolvida uma extensão do codificador HEVC voltada para a codificação de vídeos 3D, denominada 3D-HEVC. Este codificador emprega a técnica *Multi View plus Depth* (MVD) (KAUFF, 2007), que apresenta diversas vistas, onde cada uma delas é capturada por um ângulo diferente (chamadas de textura). Para cada textura há um Mapa de profundidade, que é uma representação de toda informação de profundidade dessa cena. Essas imagens são caracterizadas por possuírem bordas mais delimitadas e regiões mais homogêneas em relação as texturas. Por ser uma extensão, o 3D-HEVC parte do mesmo fundamento do HEVC e adiciona, ainda, novas funcionalidades específicas para o uso de múltiplas vistas. Como resultado, a complexidade para este codificador é ainda maior que a do HEVC.

A alta complexidade do processo de codificação traz diversos problemas para o codificador, que passa a demandar recursos de hardware e softwares mais complexos, o que pode acabar comprometendo, por exemplo, o seu uso em aplicações portáteis. Devido a isso, diversos trabalhos foram desenvolvidos no intuito de reduzir a complexidade computacional desse processo, principalmente para a codificação 2D, como em (TOHIDYPOUR, 2015). Muitos outros trabalhos também podem ser encontrados na literatura, e as aplicações para vídeos 2D podem ser aplicadas também a codificação das múltiplas vistas do 3D-HEVC. Porém, devido suas características únicas, os mapas de profundidade, que correspondem aproximadamente a 50% do tempo de codificação (SALDANHA, 2015), necessitam de soluções diferentes para reduzir sua complexidade.

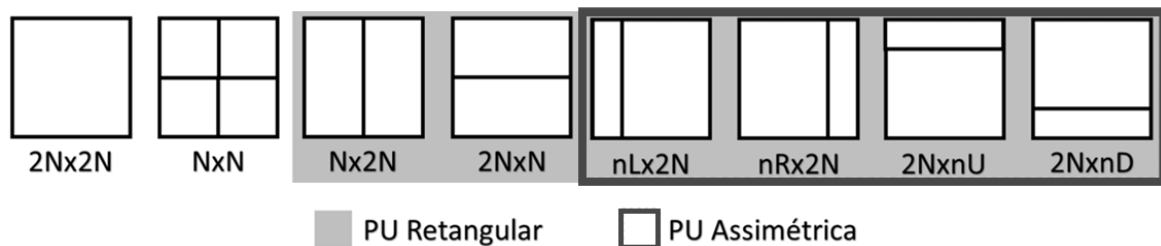


Figura 1: Modos de particionamento das PUs para os Mapas de Profundidade

É partindo desta necessidade que este trabalho propõe realizar uma análise de complexidade da codificação de profundidade do 3D-HEVC.

Um dos processos mais custosos dentro da codificação é o particionamento e subparticionamento das unidades de codificação. Este segue uma base hierárquica, onde cada quadro de um vídeo é dividido em *Coding Tree Units* (CTUs), unidades de dimensão 64x64 (pixels). Estas, por sua vez, são divididas em *Coding Units* (CUs), cujos tamanhos variam entre 64x64, 32x32, 16x16 e 8x8. Estas CUs também são finalmente divididas em *Prediction Units* (PUs), que são o foco da análise deste trabalho. Diferente das outras partições, as PUs podem apresentar formatos assimétricos, por exemplo, 2NxnU. Todos os possíveis formatos destas partições são apresentados na Figura 1.

Sabendo das características únicas dos mapas de profundidade do 3D-HEVC, este trabalho se propõe a analisar a frequência de ocorrência de cada um desses possíveis formatos de *Prediction Units* e, com base nesses resultados, analisar o impacto causado pela remoção dos modos menos frequentes na codificação de profundidade do 3D-HEVC.

2. METODOLOGIA

Inicialmente, foi realizada a análise da frequência de ocorrência de cada um dos modos apresentados na Figura 1, considerando apenas o uso de mapas de profundidade. Para esta análise foram utilizadas sete sequências (*Balloons*, *Kendo*, *Newspaper_CC*, *GT_Fly*, *Poznan_Hall2*, *Poznan_Street* e *Undo_Dancer*), todas definidas nas Condições Comuns de Teste (MULLER, 2014) com 48 quadros cada. O software de referência do 3D-HEVC foi utilizado na sua versão 16.0 (CHEN, 2015). Os resultados desta análise primária são apresentados na Figura 2. Nota-se claramente uma forte discrepância entre a ocorrência dos modos quadráticos, principalmente do modo 2Nx2N, que é escolhido como melhor opção em aproximadamente 97% das vezes, e os outros seis (Nx2N, 2NxN, 2NxnU, 2NxnD, nLx2N e nRx2N), que são escolhidos em apenas 3% das vezes.

A grande diferença entre as frequências ocorre principalmente devido à grande quantidade de regiões homogêneas dos mapas de profundidade. Estes resultados motivam a análise principal deste trabalho, que consiste em observar o impacto causado na complexidade de codificação, devido a remoção desses modos menos significativos (retangulares e assimétricos). Para realizar essa observação, duas versões alternativas do software de referência foram desenvolvidas. Estas versões, referenciadas como SR-HTM (Sem Retangulares) e SA-HTM (Sem Assimétricas), têm, respectivamente, os modos assimétricos e retangulares desativados (SR) e apenas os assimétricos (SA). Assim, em vez de o codificador buscar exaustivamente pelo melhor dos oito modos para cada PU, a busca é restringida aos dois modos quadráticos que, de acordo com os resultados da análise primária, são escolhidos (por busca exaustiva) em 97% das vezes.

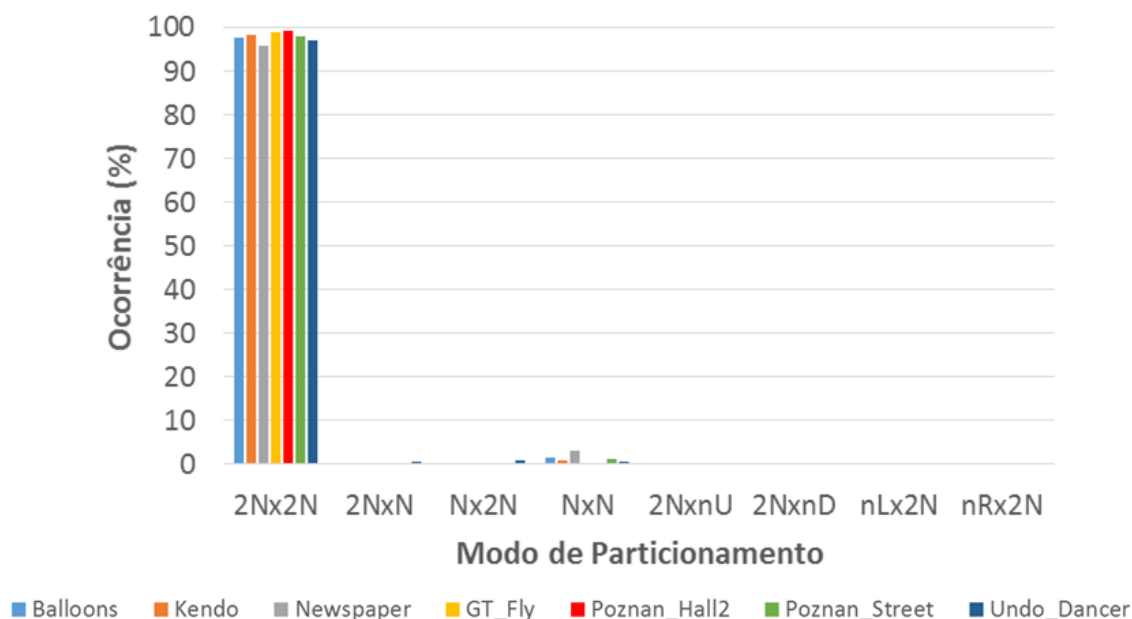


Figura 2: Ocorrências dos modos de particionamento das PUs para os Mapas de Profundidade

Assim, em vez de o codificador buscar exaustivamente pelo melhor dos oito modos para cada PU, a busca é restringida aos dois modos quadráticos que, de acordo com os resultados da análise primária, são escolhidos (por busca exaustiva) em 97% das vezes. Isso poupa boa parte do trabalho do codificador, pois reduz muito os testes necessários.

As mesmas sete sequências definidas pelas CCT foram utilizadas, cada uma com quatro diferentes conjuntos de QPs (25/34, 30/39, 35/42 e 40/45) e com o número máximo de quadros. Para fins de comparação, três amostras de dados foram simuladas. Ao todo, cada sequência foi codificada 12 vezes, resultando num total de 84 simulações considerando apenas a análise principal deste trabalho.

3. RESULTADOS E DISCUSSÃO

Na Tabela 1 são apresentados os resultados das simulações descritas. São apresentados dados tanto para o SR-HTM bem como para o SA-HTM nos diferentes QPs utilizados na codificação. A complexidade é medida em tempo de codificação. Porém, não se pode negligenciar a possível alteração na imagem que a remoção de modos pode acarretar. Por isso, também são apresentados dados de BD-Rate (BJONTEGAARD, 2001), métrica utilizada para avaliar a taxa de bits necessária mantendo a mesma qualidade de imagem (em relação aos valores obtidos pelo software de referência), para ambos os softwares.

Versão	Redução de Tempo Total	Redução de Tempo (MP)	BD-Rate
SA-HTM	0,3%	1,2%	0,041%
SR-HTM	2,6%	6,1%	0,084%

Tabela 1: Resultados de Complexidade e BD-Rate

Percebe-se, em média, uma redução de 2,6% no tempo de codificação total, lembrando que as modificações realizadas afetam apenas a codificação dos mapas de profundidade (referentes a 50% do tempo de codificação), com um custo médio de 0,084% em termos de perda na eficiência de codificação. A redução de complexidade é de 6,1% quando levamos em consideração apenas a codificação dos mapas de profundidade. É importante salientar que a perda na qualidade da imagem se refere apenas às vistas sintetizadas, pois são as únicas que utilizam os mapas de profundidade.

Os resultados encontrados em ambas as análises deste artigo estão de acordo com a teoria baseada nas características dos mapas de profundidade. Devido à grande presença de regiões homogêneas, é plausível supor que a maioria das subpartições terá formato quadrático, o que corrobora a alta discrepância encontrada entre os oito possíveis formatos de PUs na análise primária. Também de acordo com o esperado, a análise principal mostrou que, retirando seis de oito possíveis testes, há uma redução na complexidade e, já que esses seis modos são raramente escolhidos como os melhores, há pequenas perdas na eficiência de compressão.

4. CONCLUSÕES

Este trabalho apresentou uma análise de eficiência e complexidade levando em consideração a remoção dos modos de particionamento de *Prediction Units* do codificador estado da arte 3D-HEVC para os mapas de profundidade. As versões modificadas do software, com as remoções dos modos menos frequentes, atingiu em média uma redução de 6,1% na complexidade de codificação dos mapas de profundidade ao custo irrisório de 0,084% de aumento médio em BD-Rate.

Por fim, a remoção dos modos retangulares mostra-se como uma possível solução para atenuar a alta complexidade de codificação do 3D-HEVC, podendo ser utilizada em conjunto com outras soluções para redução de complexidade da codificação de textura, consequentemente facilitando diversas aplicações do codificador como, por exemplo, uma descrição de hardware.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- JCT-VC. High Efficiency Video Coding text specification draft 10, doc. JCTVC-L1003, Genebra, Suíça, 2013.
- P. Kauff; et al; "Depth map creation and image based rendering for advanced 3DTV services providing interoperability and scalability", **Signal Process., Image Commun.**, vol. 22, no.2, pp. 217–234, 2007.
- H. R. Tohidypour; et al.; "Online Learning-based Complexity Reduction Scheme for 3D-HEVC", **IEEE Transactions on Circuits and Systems for Video Technology**, vol. PP, no.99, 2015.
- Mario Saldanha; et al.; "Complexity reduction for the 3D-HEVC depth maps coding", **2015 IEEE International Symposium on Circuits and Systems**, pp. 621-624, 2015.
- K. Muller; A.Vetro; "Common Test Conditions of 3DV Core Experiments." Document: JCT3c-G1100. San José, 2014.
- Y. Chen; et al.; "Test Model 11 of 3D-HEVC and MV-HEVC", ISO/IEC JTC1/SC29/WG11, MPEG/m36142, Document: JCT3V-K1003, 2015.
- G. Bjontegaard, "Calculation of Average PSNR Differences Between RD Curves," VCEG-M33, ITU-T SG16/Q6 VCEG, 13th VCEG Meetin., 2001.